**NADYLA ALFIENA CANDRA**

**SUMMARY 4\_05\_PTYN-KS17\_012**

**MEMPELAJARI MACHINE LEARNING DENGAN PYTHON**

**Machine Learning**

Sebuah AI yang menyediakan system sebuah experience dalam sebuah coding atau menyuruh machine untuk menjalankan sebuah data tanpa harus kita memasukkan sebuah data yang diselesaikan dengan alogaritmanya. Dalam Machine Learning terdapat 2 jenis data :

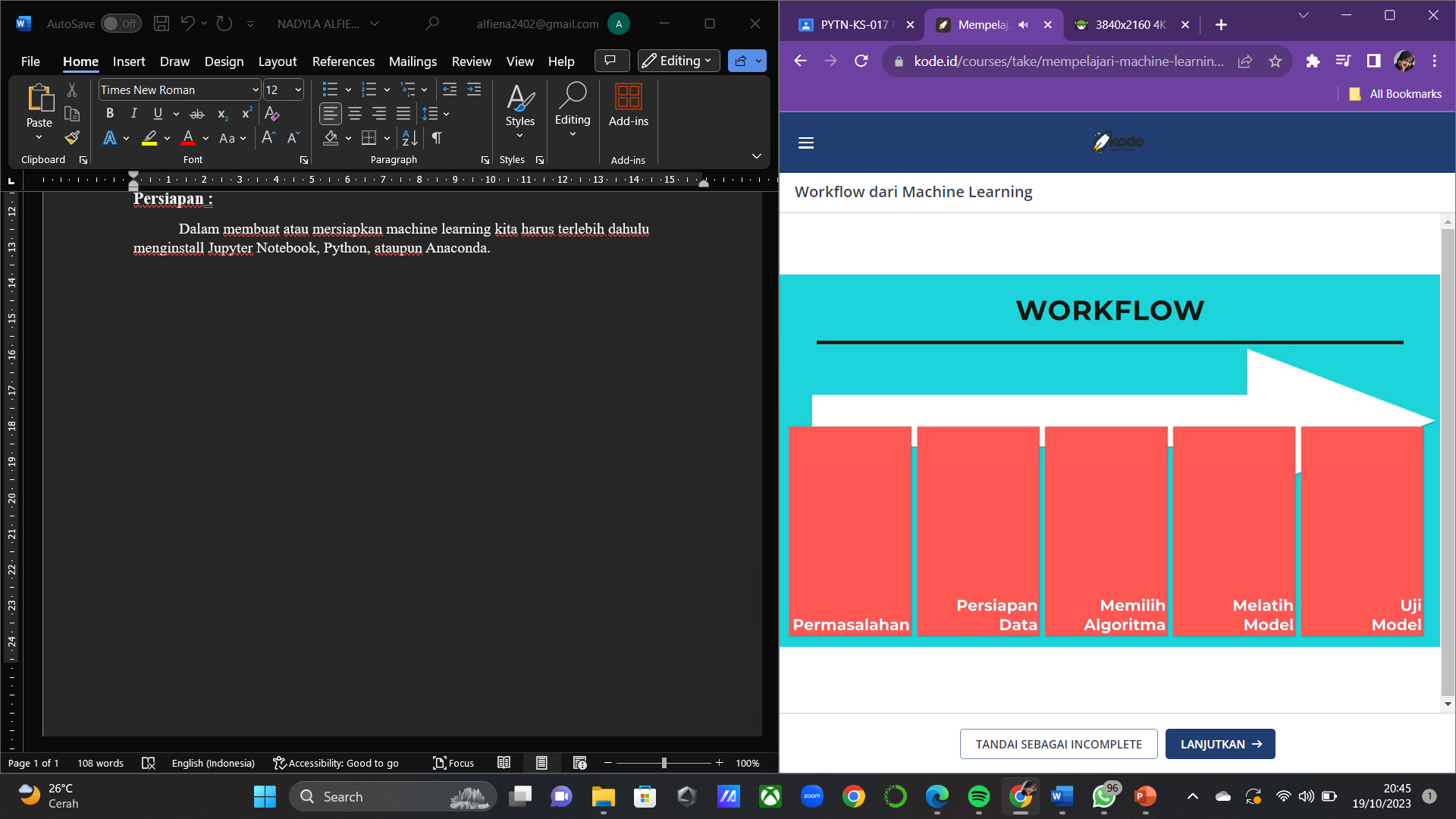
1. **Supervised**

Dari data yang nanti diolah oleh alogaritma yang nantinya kita akan mendapatkan model baru yang nantinya kita akan mendapatkan sebuah prediction. Seperti memprediksi harga rumah.

1. **Unsupervised**

Kita akan mendapatkan sebuah data cluster, mengkelompokkan data yang sama, di tiap – tiap data.

**WORKFLOW :**

****

Dalam hal ini kita harus mengidentifasi permasalahan yang dimaksud dalam sebuah kasus, contoh nya berapakah harga rumah tersebut,dsb. Selanjutnya mempersiapkan data yang benar atau akurat, selanjutnya memilih alogaritma, seperti bagaimana konsep atau alogartima mana yang cocok untuk data A atau data B ini. Setelah dirasa memilih alogaritma yang tepat untuk sebuah data maka selanjutnya kita akan melatih yang nantinya akan menguji data yang mendapatkan alogaritma yang tepat dengan sebuah model agar mendapatkan tingkat akurasi.

**Mendapatkan Data :**

Dalam membuat atau mersiapkan machine learning kita harus terlebih dahulu menginstall Jupyter Notebook, Python, ataupun Anaconda.

Tidy Data yaitu sebuah data yang lebih mudah dimanipulasi, dalam data ini harus memiliki sebuah variabel dan observasi berupa kolom dan baris. Data didapatkan dari google, data pemerintahan dan data perusahaan. Format tidak bisa sesuai dengan keinginan kita, dan kita harus load, clean dan inspect data itu sendiri.

**Memilih Alogaritma :**

Dalam memanggil alogaritma ada beberapa factor :

1. **Learning Type :**
   1. **Supervised**

Dari data yang nanti diolah oleh alogaritma yang nantinya kita akan mendapatkan model baru yang nantinya kita akan mendapatkan sebuah prediction. Seperti memprediksi harga rumah.

* 1. **Unsupervised**

Kita akan mendapatkan sebuah data cluster, mengkelompokkan data yang sama, di tiap – tiap data.

**Memilih Alogaritma berdasarkan Kasus :**

* **Naïve Bayes**
* **Logistic Regression**
* **Decision Tree**

**Melatih Model :**

* **Mengenal Model Training**

Melatih alogaritma yang sudah dipilih untuk menghasilkan model yang spesifik dan prediction model yang spesifik, karena menurut work flow apakah data sudah sesuai atau apa belum, apakah konsep sudah sudah atau benar,dsb. Menaikan performance data kita memberikan data baru.

* **Proses Training**

Setelah kita mempunyai data kita memasukkan alogaritma. Berikut merupakan proses training :

* Split Data
* Train Model
* Evaluasi Model